

# BAYESIAN ANALYSIS OF ENDOGENOUS DELAY THRESHOLD MODELS\*

GARY KOOP <sup>†</sup>

*Department of Economics, University of Glasgow,  
Glasgow, U.K. G12 8RT  
E-mail: G.Koop@socsci.gla.ac.uk*

SIMON M. POTTER <sup>‡</sup>

*Federal Reserve Bank of New York  
New York, NY 10045-0001, U.S.A.  
E-mail: Simon.Potter@ny.frb.org*

September 2000

## ABSTRACT

We develop Bayesian methods of analysis for a new class of Threshold Autoregressive models: Endogenous Delay Threshold. We apply our methods to the commonly-used sunspot data set and find strong evidence in favor of the EDTAR model over linear and traditional threshold autoregressions.

Keywords: Nonlinearity, Threshold Autoregression, Markov Chain Monte Carlo, Gibbs Sampler

---

\*We are grateful for useful comments from referees, an associate editor and the editor Ruey Tsay.

<sup>†</sup>Financial support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

<sup>‡</sup>Financial support from the NSF under grant SES 9211726 and the Center for Computable Economics at UCLA is gratefully acknowledged. The views expressed here are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System.

# 1 Introduction

In recent years there has been a growing realization of the restrictiveness of linear models in time series analysis. Accordingly, interest has increased in various nonlinear specifications. Tong (1990) provides an introduction and discussion of many commonly-used nonlinear models. There are two related problems with nonlinear models:

1. They can be very parameter rich, and thus be subject to the criticism that they are merely mining the data.
2. Theory often does not tell us which nonlinear functional form is appropriate. That is, it is easy to say that linear models are too restrictive, it is much harder to choose one of the infinite number of departures from linearity that are possible.

In this paper we consider Bayesian estimation for a particular class of models called Endogenous Delay Threshold Autoregressive (EDTAR) models. The goal is to alleviate the two problems mentioned above. In particular EDTAR models represent a highly flexible class of nonlinear time series models yet the flexibility is not produced by over-parameterization. A related paper is Pesaran and Potter (1997) which uses classical econometric methods to analyze an EDTAR model which is a special case of the one considered in the present paper. The present approach allows for a more general and coherent treatment of a whole class of EDTAR models (as opposed to a particular variant suitable for use in a specific application) and, as we shall argue below, the adoption of a Bayesian approach simplifies some econometric problems and allows for easy comparison of multiple, possibly non-nested, models.

The EDTAR class is included in the even more general class of threshold autoregressions introduced by Tong (see Tong 1990 for a review). The threshold class is distinguished by the use of piecewise linear autoregressive models. EDTAR models represent a new implementation of the threshold approach. Most implementations of the threshold principle have used a simple classification of the time series into two regimes based on a single lag of the time series (i.e. self-exciting threshold autoregressive, or SETAR, models). This choice is convenient because, when more general regime classification or multiple regimes are considered within the standard approach, parsimony problems and estimation difficulties arise (see, e.g., Chen, 1990). EDTAR models differ from the standard implementation by using previously unexploited information about the length of time spent in regimes. This allows the construction of ‘sub-regimes’ within ‘major’ regimes. Parsimony is maintained by tightly restricting parameters across the sub-regimes.

In previous work (Koop and Potter, 1999a, 1999b), we have argued for the advantages of Bayesian methods when working with nonlinear time series models. A chief advantage of Bayesian methods is the mitigation of data-mining/overfitting problems through model averaging. Furthermore, Bayesian methods are particularly attractive for threshold models because they are capable of producing information about parameter uncertainty for the regime-classifying and AR parameters jointly. Classical techniques are only capable of producing measures of parameter uncertainty across samples conditional on the regime classification because of the faster than  $\sqrt{T}$  convergence of the regime classification parameters (Chan, 1993).

In this paper we develop Bayesian methods of analysis for EDTAR models based on Markov Chain Monte Carlo techniques. In addition we also describe a simple method for calculating Bayes factors for EDTAR models versus linear models and SETAR models based on the work of Koop and Potter (1999a). These methods are applied to the sunspot data series contained in the appendix of Tong (1990). We find considerable posterior evidence in favor of the EDTAR model over SETAR and linear models.

The remainder of the paper is organized as follows: Section 2 describes and motivates EDTAR models. Section 3 develops the Bayesian statistical techniques used in this paper. Section 4 presents estimates of an EDTAR model for Wolfe’s sunspot data and comparisons to linear and traditional threshold models based on marginal likelihoods. Section 5 concludes with a review of the main results of the paper and a discussion of extensions.

## 2 EDTAR Models

The EDTAR model developed here, and used in our empirical work, can be written as:

$$Y_t = \phi' \mathbf{X}_t + \theta' \mathbf{Z}_t(\gamma) + h_t(\gamma, \sigma) V_t, \quad (1)$$

where

$$\mathbf{X}_t = (1, Y_{t-1}, \dots, Y_{t-p})', \mathbf{Z}_t(\gamma) = (Z_{1t}(\gamma), Z_{2t}(\gamma))',$$

$$Z_{mt}(\gamma) = \mathbf{1}[I_{mt}(\gamma) > 0] (f_m(Y_t, \dots, Y_{t-s}) + Z_{mt}(\gamma)),$$

$$h_t(\gamma, \sigma) = \sigma_0 \mathbf{1}[I_{0t}(\gamma) = 1] + \sigma_2 \mathbf{1}[I_{1t}(\gamma) > 0] + \sigma_3 \mathbf{1}[I_{2t}(\gamma) > 0].$$

Exact details are given below, suffice it to note here that  $\gamma$  is a vector of parameters used for defining regimes,  $I_{mt}(\cdot)$  for  $m = 0, 1, 2$  indicate three major regimes and  $\sigma_j$  for  $m = 0, 1, 2$  are error

scalings for these regimes.  $\mathbf{Z}_t(\gamma)$  and  $f_m(\cdot)$  are defined below. The crucial ingredient in the EDTAR model is  $\mathbf{Z}_t(\gamma)$  which contains measures of overheating and cooling (e.g.  $f_m(\cdot)$  can be chosen so that  $\mathbf{Z}_t(\gamma)$  captures ceiling/floor effects as in Pesaran and Potter, 1997) of the sort that are thought to exist in many time series. We show how such a model can be written as a threshold model with the property that sub-regimes exist within three major regimes. However, the differences across regimes and sub-regimes are restricted in a natural fashion. These restrictions allow us to estimate models with a large number of regimes.

In the remainder of this section, we relate the EDTAR model to some implementations of the threshold model which have appeared in the literature. We gradually motivate and build up the components of the EDTAR model given in Equation 1. In other words, Equation 1 should be treated as the ending point of our derivations, not the beginning. We present it here only to provide a concrete view of where the following discussion is headed.

## 2.1 Background and Motivation

When one leaves the world of linear models, there are many nonlinear forms to choose from. For many types of applications, it is natural to assume that different regimes exist and, within each regime, it may make sense to assume a linear structure. For example, economic theory often tells us that the dynamics of real GDP growth should be different in expansionary times than contractionary times. Once the series is divided into regimes, the question arises as to what sort of dynamics should apply in each regime. These two problems – regime classification and selection of dynamics in each regime – lie at the heart of specification issues with threshold models.

In order to motivate the class of EDTAR models we start from the most general specification of the TAR class first given in Tong and Lim (1980). Let  $\{Y_t : t = 0, 1, \dots\}$  be a time series and let  $M_t$  be an indicator random variable taking values in the set  $\{1, 2, \dots, M\}$ . Then the canonical threshold autoregression is defined by:

$$Y_t = \alpha^{\{M_t\}} + \phi^{\{M_t\}}(L)Y_{t-1} + \sigma^{\{M_t\}}V_t, \quad (2)$$

where  $V_t$  is an IID sequence of standardized random errors with zero means and unit variances; and for  $M_t = m$ ,  $\alpha^{\{m\}}$  is a constant,  $\sigma^{\{m\}}$  are regime specific scalings of the errors and  $\phi^{\{m\}}(L)$  is a finite order polynomial in the lag operator  $L$ . In general, the order of the polynomial as well as the coefficients can vary with the index  $M_t$ .

Most applications of threshold autoregressions have simplified the problem of choice of the index variable by choosing a lag of the observed time series  $d$ , known as the delay parameter, and a collection of thresholds  $\{r_m, m = 1, \dots, M-1\}$  to construct  $M_t$ . An example would be:

$$Y_t = \begin{cases} \alpha_1 + \phi_1(L)Y_{t-1} + \sigma_1 V_t & \text{if } Y_{t-d} < r_1, \\ \alpha_2 + \phi_2(L)Y_{t-1} + \sigma_2 V_t & \text{if } r_1 \leq Y_{t-d} < r_2, \\ \vdots & \vdots \\ \alpha_K + \phi_M(L)Y_{t-1} + \sigma_K V_t & \text{if } Y_{t-d} > r_{M-1}. \end{cases} \quad (3)$$

This model is known as a self-exciting threshold autoregressive model, or SETAR. Since the index variable is constructed from the location of observable lags of the time series, the specification is relatively easy to estimate, test and evaluate as exemplified in Tong (1990). A common alternative class of models, based on one originally presented in Hamilton (1989), has  $M_t$  determined by a hidden Markov chain.

In this paper we develop a different version of the threshold model by introducing restrictions across regimes. These restrictions allow us to estimate models with a large number of regimes. As an illustration, we give the form of the threshold model, equivalent to equation (1), which we estimate later for the sunspot data:

$$Y_t = \begin{cases} \alpha + \phi_p(L)Y_{t-1} + \sigma_0 V_t & \text{if } r_2 \leq \Delta Y_{t-1} \leq r_1, \\ \alpha + \phi_p(L)Y_{t-1} + \theta_1(Y_{t-1} - Y_{t-2} - r_1) + \sigma_1 V_t & \text{if } \Delta Y_{t-1} > r_1 \\ \vdots & \vdots \\ \alpha + \phi_p(L)Y_{t-1} + \theta_1 \sum_{s=1}^j (Y_{t-s} - Y_{t-s-1} - r_1) + \sigma_1 V_t & \text{if } \prod_{s=1}^j 1(\Delta Y_{t-s} > r_1) > 0, \\ \vdots & \vdots \\ \alpha + \phi_p(L)Y_{t-1} + \theta_2(Y_{t-1} - Y_{t-2} - r_2) + \sigma_2 V_t & \text{if } \Delta Y_{t-1} < r_2 \\ \vdots & \vdots \\ \alpha + \phi_p(L)Y_{t-1} + \theta_2 \sum_{s=1}^j (Y_{t-s} - Y_{t-s-1} - r_2) + \sigma_2 V_t & \text{if } \prod_{s=1}^j 1(\Delta Y_{t-s} < r_2) > 0, \\ \vdots & \vdots \end{cases} \quad (4)$$

where  $\Delta Y_t = Y_t - Y_{t-1}$ . Note that the number of regimes actually visited by the time series is endogenous and depends on how many consecutive periods the change in the sunspot numbers are above or below a threshold. Further, rather than adding  $p+3$  parameters for each extra regime (as a SETAR would do), no extra parameters are added as the change in the sunspots stays above or below the threshold for more than one consecutive period. Instead the conditional expectation function adapts through the change in the number of lags of the difference in the sunspots included.

To return to the two central problems of threshold modeling described above, it can be seen

that the standard SETAR model given above solves them in a very simple manner. Regimes are classified on the basis of observable lags of the series, and the dynamics in each regime are allowed to follow a totally different autoregressive process. However, this strategy can be unsatisfactory if the number of regimes is large. In the example of real GDP dynamics, one might expect different regimes to apply in the case where the economy has been in a recession, expansion or normal times. Further, within each of these major regimes one might expect sub-regimes to exist. For instance, if the economy has been in a recession a short time different dynamics might apply relative to a case where a long recession has occurred. Hence, any time series model of the business cycle could end up with a large number of regimes. If the dynamics in each regime are left unrestricted, problems of parsimony quickly arise.

This tension between wanting flexible models with many regimes and the parsimony problems which result, is reflected in much of the recent literature. In our discussion of some of this literature, we will let  $X_t$  = the log of US real output and  $Y_t = X_t - X_{t-1}$ . Consider the paper of Tiao and Tsay (1994), which uses a SETAR model in an application involving US real GNP growth. The authors address the issue of regime classification by beginning with a model identical to Equation 3 above with two regimes. They then search over possible values of  $d$  and  $r_1$  and select  $d = 2$  and  $r_1 = 0$ , values which provide the most evidence of nonlinearity. They then retain these values when they expand to a four regime model with regimes defined as:

- Regime I:  $Y_{t-1} \leq Y_{t-2} \leq 0$ . (i.e., a worsening recession).
- Regime II:  $Y_{t-1} > Y_{t-2}$  but  $Y_{t-2} \leq 0$  (i.e., recovery from a recession).
- Regime III:  $Y_{t-1} \leq Y_{t-2}$  but  $Y_{t-2} > 0$  (i.e., slowdown in growth/beginning of a recession).
- Regime IV:  $Y_{t-1} > Y_{t-2} > 0$  (i.e., a strengthening economy).

Statistical evidence is provided on the desirable properties of this model. The Tiao and Tsay paper illustrates some important issues which arise with SETAR modelling. First, a careful statistical analysis of many macroeconomic time series (in their case, real GNP growth) indicates evidence for many regimes (in their case, 4). Second, SETAR models tend to have many parameters in them. Tiao and Tsay keep the number of parameters manageable by sticking with regimes based on  $d = 2$  and threshold values of zero, but even their model has 13 parameters (i.e. a different intercept and error variance for each regime plus they model three regimes as AR(1) and the fourth AR(2)). Third, there are so many choices to be made in even a simple

model (e.g. number of regimes, definition of regimes, delay and threshold parameters in each regime, lag lengths in each regime, etc.) that data mining can be a problem. Fourth, even with a parameter-rich model such as Taio and Tsay's there are always possible criticisms of particular aspects of the model (e.g. why focus only on the selected threshold and delay parameters? Why not consider other regime definitions?, etc.).

In contrast to the SETAR tradition, there is another class of specifications which can be illustrated using the model of Beaudry and Koop (1993). In an application involving US real GNP, Beaudry and Koop add lags of a variable they call the current depth of recession (*CDR*) to a standard ARMA(p,q) model. This variable is defined as the gap between the current level of real GNP and its historical maximum **including the current value**. That is,

$$CDR_t = X_t - \max\{X_t, X_{t-1}, \dots, X_0\}.$$

They justified the inclusion of this variable on the grounds that it would allow for shocks to real GNP to have asymmetric effects (e.g. since technical regress is unlikely negative shocks to real GNP would not have a permanent effect, but positive shocks could have a permanent effect). Beaudry and Koop present evidence that such *CDR* models were statistically superior to linear ARMA models, at least for the case of US real GNP. Hess and Iwata (1997) find the *CDR* variable to be significant in the case of US GDP, but not in the case of GDP for some other countries. In a comparison involving several different nonlinear and linear specifications, Jansen and Oh (1999) find the *CDR* model to be the best model for the case of US real GNP.

The *CDR* model has the advantages that it is quite parsimonious (i.e., it only adds  $q$  extra parameters to a linear AR specification, where  $q$  is the number of lags of the *CDR* variable included in the model), based in economic theory and seems to provide a good data description of some macroeconomic variables such as real GNP. However, its disadvantage is that it was developed in light of theories of behavior relating particularly to real output, and it is probably too restrictive to be considered as a class of nonlinear models that is suitable for use with a wide variety of time series.

Pesaran and Potter (1997) point out the relationship between the *CDR* model and TAR class and show that the model of Beaudry and Koop (1993) is, in fact, somewhat similar to that of Tiao and Tsay (1994). The model of Beaudry and Koop is:

$$Y_t = \alpha + \phi_p(L)Y_{t-1} + \theta CDR_{t-1} + V_t. \tag{5}$$

It can be verified that an alternative way of writing the  $CDR$  variable is provided in the following two equations (see Pesaran and Potter, 1997, for additional details):

$$F_t = \mathbf{1}(CDR_{t-1} + Y_t < 0), \quad t = 1, 2, \dots, \quad (6)$$

$$CDR_t = (CDR_{t-1} + Y_t) F_t, \quad t = 1, 2, \dots, \quad (7)$$

where  $CDR_0 = 0$  and  $\mathbf{1}(\cdot)$  is the indicator function equal to 1 if the statement inside the parenthesis is true and 0 otherwise.  $F_t$  is, thus, an indicator function for what Pesaran and Potter call the “floor” regime.

The connection between the CDR and TAR models can be found by expanding out Equation 7 for  $j$  periods:

$$CDR_t = (CDR_{t-1} + Y_t) F_t, \quad CDR_t = (CDR_{t-2} + Y_{t-1}) F_{t-1} F_t + Y_t F_t,$$

and so on, until:

$$CDR_t = \sum_{i=0}^j Y_{t-i} \prod_{\tau=0}^i F_{t-\tau} + CDR_{t-j-1} \prod_{\tau=0}^j F_{t-\tau}.$$

By construction, it is possible for this recursion to continue into the infinite past. However, if the time series  $\{Y_t\}$  is stationary then the maximum lag will be finite (with probability one). In the  $CDR$  model a sufficient condition for stationarity is that  $-1 < \theta < 0$ , and the polynomial  $1 - \phi_p(z)$  has all its roots outside the unit circle. In this case the  $CDR_t$  variable will have the effect of dampening the (negative) fluctuations in output. Since  $Y_t$  is the first difference of output, this dampening effect is similar to a reflective barrier or a floor determined by the previous maximum level of output. Thus, the index variable,  $F_t$ , switches on when current output is below the maximum level of output achieved (i.e., is in the floor regime).

Pesaran and Potter also show that the above recursion can be used to rewrite the model in Equation 5 as a member of the TAR class:

$$Y_t = \begin{cases} \alpha + \phi_p(L)Y_{t-1} + V_t & \text{if } F_{t-1} = 0, \\ \alpha + \phi_p(L)Y_{t-1} + \theta Y_{t-1} + V_t & \text{if } F_{t-1} = 1 \text{ and } F_{t-2} = 0, \\ \vdots & \vdots \\ \alpha + \phi_p(L)Y_{t-1} + \theta \sum_{s=1}^j Y_{t-s} + V_t & \text{if } \prod_{s=1}^j F_{t-s} = 1 \text{ and } F_{t-j-1} = 0 \\ \vdots & \vdots \end{cases} \quad (8)$$

Note that Equation 8 implies that the  $CDR$  model has many sub-regimes within the floor regime. In other words, dynamic properties can differ from the period when the economy just enters a



recession, to when it has been in a recession for 1 period, for 2 periods, etc. Hence, it is similar (but not identical to) the multiple-regime model of Tiao and Tsay (1994). However, the *CDR* model is much more parsimonious. The differences between dynamics in the sub-regimes depends only on one parameter,  $\theta$ . In other words, problems of parsimony in the multiple-regime TAR model are overcome by restricting the dynamics in the sub-regimes to be similar to one another. The restrictions fall out naturally from an intuitively appealing concept: that the depth of a recession should affect real output dynamics.

Pesaran and Potter (1997) extend the basic *CDR* variable to allow for an unknown threshold parameter (i.e. instead of measuring accumulated drops in real GDP, it measures accumulated drops below a threshold). Then they add a ceiling regime. The ceiling regime is defined as the opposite of the floor regime. The variable on which it is based reflects overheating, *OH*, and measures accumulated growth above a threshold. Pesaran and Potter (1997) derive classical estimation and testing methods for the floor and ceiling model and find strong statistical evidence in its support for the case of real US GDP.

The floor and ceiling model of Pesaran and Potter takes a threshold form where the number of regimes is determined by the data. Hence, this model is sometimes referred to as an endogenous delay threshold autoregressive or EDTAR model (a terminology which we adopt in the present paper). Although it is quite flexible, the specification used by Pesaran and Potter was designed particularly to model real GDP. There are several modelling choices (e.g. precise definitions of *CDR* and *OH* and the decision to use only one lag of these variables) that may be questionable in other empirical contexts. Hence, a chief aim of the present paper is to generalize the EDTAR model used in Pesaran and Potter (1997) into a whole class of models with similar properties. We believe that such generalizations might be useful in a wide range of fields of study. After all, ideas relating to reflective barriers, dampening of shocks, overheating and undercooling, etc. appear in many areas of science and social science. Hence, a general variant of the EDTAR model is developed in the following sub-sections. In addition to using the floor and ceiling model to provide intuition about the workings of our more general model we also discuss the model used in the application to the sunspot data in Section 4.

## 2.2 Endogenous Delay Index Variables

The first issue we address is how to define the different regimes in a flexible way. Consider the manner in which the TAR model in Equation 3 generates regimes for a fixed delay variable. It

uses the value of  $Y_{t-d}$  to split up the time series up into separate regimes. One weakness of this approach is that there is no good reason to isolate one particular lag of the time series to define the regimes. One natural way of avoiding the need to fix a delay variable is to allow information from previous values of the time series to feed back into the construction of the variable used to define a particular regime. For example, consider a regime which is activated by large values of the time series (e.g. an expansionary regime where fast growth is occurring). Instead of choosing one lag to measure large values, one could switch on an index variable when the time series crosses a particular threshold and then measure how long the time series spent within this regime. The activation of the ‘major’ (e.g. expansionary) regime leads to sub-regimes indexed by the time spent in that regime (e.g.  $x$  quarters spent in an expansion). The idea being that the behavior of the time series will be different if, for example, it has only spent 1 period in the regime *versus* 10 periods.

In order to formally develop these ideas we concentrate on the case of three major regimes. As discussed above this is a reasonable choice in many circumstances (e.g. the idea of floor and ceiling reflective barriers is sensible in many cases), but more major regimes could be added at little cost. The index variables required to produce three major regimes, which we call endogenous delay variables  $\{I_{mt}, m = 0, 1, 2\}$ , are generated by a feedback relationship amongst past values of the time series. The feedback relationship has a particular recursive ordering controlled by a parameter,  $\varsigma$ , that takes on the value 0 or 1. The role of this parameter is to ensure that two regimes cannot occur at the same time. In the context of real GNP work, it is probably reasonable for the researcher to set  $\varsigma = 0$ . This would ensure that “rapid recovery from a recession” is not labelled “overheating”. However, in other applications it may not be more sensible to set  $\varsigma$  at 0 or 1 based on some *a priori* knowledge. Accordingly, in the present case we treat  $\varsigma$  as an unknown parameter which is estimated from the data.

$$I_{1t} = ((1 - \varsigma)\mathbf{1}[I_{2t} = 0] + \varsigma) \sum_{j=J_1}^t \prod_{s=1}^j \mathbf{1}[f_1(Y_{t-s}, \dots, Y_{t-s-K_1}) > 0], \quad (9)$$

$$I_{2t} = (\varsigma\mathbf{1}[I_{1t} = 0] + (1 - \varsigma)) \sum_{j=J_2}^t \prod_{s=1}^j \mathbf{1}[f_2(Y_{t-s}, \dots, Y_{t-s-K_2}) < 0], \quad (10)$$

$$I_{0t} = \mathbf{1}[I_{1t} + I_{2t} = 0] + I_{0t-1} \quad (11)$$

where the functions  $f_m, m = 1, 2$  are given by

$$f_m(Y_{t-s}, \dots, Y_{t-s-K_m}) = \begin{cases} Y_{t-s} - r_m & \text{if } K_m = 0 \\ \frac{Y_{t-s} - Y_{t-s-K_m}}{K_m} - r_m & \text{if } K_m > 0 \end{cases} \quad (12)$$

where  $K_m, m = 1, 2$  are nonnegative integers,  $J_m, m = 1, 2$  is an integer greater than or equal to 1.

The endogenous delay index variable indicates the major regime and sub-regimes (i.e. if  $I_{mt} > 0$  we say major regime  $m$  is occurring, and  $I_{mt} = 1, 2, 3, \dots$  indicates various sub-regimes). To simplify the exposition we assume that  $t > \max_m(J_m + K_m)$ .

Note that we now have three major regimes and index variables ( $I_{mt}$ ) which count the number of periods the time series has been in each regime. However, the precise definition of the regimes depends on the parameters  $K_m, J_m$  and  $r_m$ . It is worthwhile to provide additional intuition for the precise role that these parameters play in defining the regimes. In the models discussed above for real output, the major regimes were defined based on lags of the time series. Since  $Y_t$  was real output growth this was sensible. That is, regimes were defined using concepts like “real output is falling” or “real output is growing fast”. But, in other applications, it might make sense for regimes to be defined in other ways. In the present general class of EDTAR models, we allow for regimes to be defined in the traditional manner, based on past *levels* of the series ( $K_m = 0$ ) or on past *changes* on the series ( $K_m = 1$ ) and past *average changes* over the last  $K_m$  periods ( $K_m \geq 2$ ). We treat  $K_m$  as an unknown parameter which can be estimated from the data. Hence, we let the data determine which the most appropriate regime definitions are.

The parameters  $J_m$  allow a further degree of flexibility in defining the regimes. This allows for delays in regime activation. For instance, the floor and ceiling model of Pesaran and Potter (1997) implied that the series entered the floor regime the period after a recession hit. But (especially given the  $d = 2$  findings of Potter, 1995 and Tiao and Tsay, 1994 using SETAR models) it might be more sensible to allow the regime switch to occur two periods after the recession hits. By treating  $J_m$  as an unknown parameter to be estimated, we let the data decide how long a delay must occur before a regime switch is activated.

To preview our empirical findings, consider the sunspot model with the most posterior support from Section 4. This has  $J_m = K_m = 1, m = 1, 2$ . Thus, the two major regimes 1, 2 are entered on the first activation of the indicator function:  $(Y_{t-1} - Y_{t-2}) > r_1$ , or  $(Y_{t-1} - Y_{t-2}) < r_2$ . For the sunspot data, these regimes are mutually exclusive and the parameter  $\varsigma$  is redundant. The time series stays in each major regime for as long as the change in the number of sunspots exceeds the threshold. As soon as the change in the number of sunspots falls within  $[r_2, r_1]$  the

model returns to regime 0.

### 2.3 Parameter restrictions across sub-regimes

Now that we have defined all the major regimes and sub-regimes, we must decide how to use this information. In Pesaran and Potter (1997), one lag of the *CDR* and *OH* variables are added to an AR(p) specification. In the present paper, we do something similar but more general.

If all the sub-regime information is used to construct a conventional TAR model as in Equation 3 the number of parameters would be enormous. In order to reduce the number of parameters, we start by restricting the error variance to change only across major regimes. Next, for the major regime defined by  $I_{0t} > 0$ , we assume that the coefficients do not change across the sub-regimes. In other words, in “normal times” dynamics do not change. Hence, we model this major regime in a conventional manner with an autoregression of order  $p$ . We have

$$Y_t = \alpha + \phi_p(L)Y_{t-1} + \sigma_0 V_t, \text{ if } I_{0t} > 0.$$

Now consider the activation of one of the other two major regimes. In these cases, we include the functions,  $f_m(\cdot)$  given in equation (12) as additional explanatory variables.

Consider first the case where  $K_m = 0$ . For the  $j^{th}$  period of activation of major regime  $m$ , we use the following specification which allows the first  $j$  lags of the time series and intercept to change:

$$Y_t = \alpha + \phi_p(L)Y_{t-1} + \sum_{s=1}^j \xi_{ms}(Y_{t-s} - r_m) + \sigma_m V_t, \text{ if } I_{mt} = j. \quad (13)$$

To return to the real GDP example, when the economy leaves normal times and enters an expansion, at first (e.g.  $I_{1t} = 1$ ) the dynamics only change slightly (i.e. only the intercept and AR(1) coefficient change). As the expansion continues longer, the dynamics can differ more and more. If the expansion were to continue for  $p$  periods, then the AR representation would be totally unrestricted relative to that in normal times.

In the case where  $K_m > 0$ , the dynamic model if regime  $m$  is activated ( $m = 1, 2$ ) is:

$$Y_t = \alpha + \phi_p(L)Y_{t-1} + \sum_{s=1}^j \xi_{ms} \left( \frac{Y_{t-s} - Y_{t-s-K_m}}{K_m} - r_m \right) + \sigma_m V_t, \text{ if } I_{mt} = j. \quad (14)$$

This completely defines the most general EDTAR model discussed in this paper. Note that, if we knew what  $K_m$ ,  $J_m$  and  $r_m$  were, then we could calculate  $I_{mt}$  and the model would be in the form of a linear regression model with heteroskedasticity (i.e. the error variance changes

across the major regimes). This indicates that the model is relatively simple to interpret and work with (see discussion of estimation and inference below).

In order to provide some additional intuition, we now focus on the best sunspot model where  $J_m = K_m = 1, m = 1, 2$ . Extensions to other values are straightforward. We can write the EDTAR model as:

$$Y_t = \begin{cases} \alpha + \phi_p(L)Y_{t-1} + \sigma_0 V_t & \text{if } I_{0t} > 0, \\ \alpha + \phi_p(L)Y_{t-1} + \xi_{11}(Y_{t-1} - Y_{t-2} - r_1) + \sigma_1 V_t & \text{if } I_{1t} = 1 \\ \vdots & \vdots \\ \alpha + \phi_p(L)Y_{t-1} + \sum_{s=1}^j \xi_{1s}(Y_{t-s} - Y_{t-s-1} - r_1) + \sigma_1 V_t & \text{if } I_{1t} = j, \\ \vdots & \vdots \\ \alpha + \phi_p(L)Y_{t-1} + \xi_{21}(Y_{t-1} - Y_{t-2} - r_2) + \sigma_2 V_t & \text{if } I_{2t} = 1 \\ \vdots & \vdots \\ \alpha + \phi_p(L)Y_{t-1} + \sum_{s=1}^j \xi_{2s}(Y_{t-s} - Y_{t-s-1} - r_2) + \sigma_2 V_t & \text{if } I_{2t} = j, \\ \vdots & \vdots \end{cases} \quad (15)$$

Equation (15) illustrates that the EDTAR model still belongs in the general TAR class.

In practice this general EDTAR model is over-parameterized. Hence, we consider a tight restriction on the  $\xi_{ms}$  coefficients for each regime which we have found to be reasonable and which relates to the floor and ceiling model of Pesaran and Potter. To motivate this restriction, consider how the dynamics change as the time series leaves the base regime 0. For the case  $K_m = 1$ , the total change in the conditional mean from  $I_{0t-j} > 0$  to  $I_{mt} = j$  due to the threshold structure is:

$$Z_{mt}^* = \sum_{s=1}^j \xi_{ms}(Y_{t-s} - Y_{t-s-1} - r_m).$$

In other words, if the series has spent  $j$  periods in regime  $m$ , the deviation from the basic AR(p) model is given by  $Z_{mt}^*$ .

For a given  $I_{mt}$ ,  $Z_{mt}^*$  is a moving average of the past values of the time series,  $\{Y_{t-s} - Y_{t-s-1} - r_m\}$ . The floor and ceiling model of Pesaran and Potter implicitly imposed the restriction that each term in this moving average receives equal weight. In other words, it is the simple sum that matters (e.g. the accumulated drops in the sunspots below a threshold). We also impose this restriction in our modelling. That is, we let

$$Z_{mt}^* = \theta_m \sum_{s=1}^j (Y_{t-s} - Y_{t-s-1} - r_m),$$

i.e.,  $\xi_{m1} = \xi_{m2} = \dots = \xi_{ms} = \dots$ . This restricts the formulation of equation (15) to that found in equation (4) of Section 2.1. Further, it means that one can write the  $Z_{mt}$  variables in the convenient recursive form:

$$Z_{mt} = 1(I_{mt} > 0) \{f_m(Y_t, \dots, Y_{t-s}) + Z_{mt-1}\}.$$

Hence, in the following material we work with the EDTAR model set out previously but restrict equations (13) and (14) to:

$$Y_t = \alpha + \phi_p(L)Y_{t-1} + \theta_m Z_{mt} + \sigma_m V_t, \text{ if } I_{mt} = j. \quad (16)$$

With the sunspot data, we find most empirical support for  $J_m = K_m = 1, m = 1, 2$ , which implies  $Z_{mt} = \sum_{s=1}^j (Y_{t-s} - Y_{t-s-1} - r_m)$ .

For future reference, note that linearity in the conditional mean is present if  $\theta_1 = \theta_2 = 0$ . Homoskedasticity is present if  $\sigma_0 = \sigma_1 = \sigma_2$ .

### 3 Estimation and Inference

Bayesian techniques for analyzing SETAR models are given in Geweke and Terui (1993) and Chen and Lee (1995). Geweke and Terui derive the posterior density of the parameters and suggest the use of Monte Carlo integration for drawing posterior and predictive inferences. Unfortunately, similar Monte Carlo integration techniques do not extend to the EDTAR case because of the parameter restrictions across regimes and the presence of the threshold value in the construction of  $Z_{mt}$ . Instead we use Markov Chain Monte Carlo (MCMC) techniques in a similar manner to Chen and Lee (1995). A major difference to the traditional threshold autoregression case analyzed by those authors is that rather than only having to sort the data once for each different value of the delay  $d$ , in the EDTAR model different values of the delay parameters,  $K_m, J_m, \varsigma$  and the thresholds requiring sorting the data anew for each iteration of the posterior sampler.

#### 3.1 Notation

For estimation it is convenient to write the EDTAR model in a more compact form and make explicit the dependence on the parameters  $(r_m, K_m, J_m : m = 1, 2)$ .

$$Y_t = \phi' \mathbf{X}_t + \theta' \mathbf{Z}_t(\gamma) + h_t(\gamma, \sigma) V_t = \beta' \mathbf{W}_t(\gamma) + h_t(\gamma, \sigma) V_t, \quad (17)$$

where

$$\mathbf{X}_t = (1, Y_{t-1}, \dots, Y_{t-p})', \mathbf{Z}_t(\gamma) = (Z_{1t}(\gamma), Z_{2t}(\gamma))', \mathbf{W}_t(\gamma) = (\mathbf{X}_t', \mathbf{Z}_t(\gamma)')',$$

$$h_t(\gamma, \sigma) = \sigma_0 \mathbf{1}[I_{0t}(\gamma) = 1] + \sigma_2 \mathbf{1}[I_{1t}(\gamma) > 0] + \sigma_3 \mathbf{1}[I_{2t}(\gamma) > 0],$$

and

$$\begin{aligned} \gamma_m &= (r_m, \kappa_m)' \quad \gamma = (\gamma_1, \gamma_2, \varsigma)' \\ \kappa_m &= (K_m, J_m)', \quad \kappa = (\kappa'_1, \kappa'_2)', \theta = (\theta_1, \theta_2)', \\ \sigma &= (\sigma_0, \sigma_1, \sigma_2)' \\ \beta &= (\phi', \theta')', \\ \phi &= (\alpha, \phi_1, \dots, \phi_p)'. \end{aligned}$$

In addition we will also use the notation  $\mathbf{Y} = (Y_1, Y_2, \dots, Y_N)$ , for the observed sample of observations and  $\psi$  to refer to all the parameters in the model, i.e.  $(\beta, \gamma, \sigma)$ . Further we will assume that  $\gamma \in \Gamma$ , a compact set.

### 3.2 Priors

We assume that  $J_1, J_2$  are uniformly distributed, independently of each other, over the integers  $1, 2, \dots, \bar{J}$  and  $K_1, K_2$  are uniformly distributed independently of each other over the integers  $0, 1, \dots, \bar{K}$ . Further  $K_1, J_1, K_2, J_2$  are mutually independent.

Conditional on the values of  $K_1, J_1$  and assuming  $\varsigma = 1$  we construct prior distributions for  $r_1, r_2$  such that at least 15% of the data is in each of the three major regimes. Ideally, one would like to generate the 15% in each major regime using the properties of all parameters. However, since there is no closed form for the stationary distribution, this is not possible. Instead, we use the observed data. For the threshold  $r_1$ , we assume a prior density that is uniform from the smallest value that implies that 70% of the data is in major regime 1 and the largest value that implies 15% of the data is in major regime 1. Conditional on the values of  $r_1, K_1, J_1$  and  $K_2, J_2$ , the prior distribution of  $r_2$  is uniform from the smallest value that implies 15% of the data is in regime 2, to that which would imply 15% of the data in regime 0 given the values of  $I_{1t}$ . We denote the joint distribution by  $f(r_2|\kappa, r_1)f(r_1|\kappa)$ . Notice that the prior is constructed imposing a particular value of  $\varsigma$ , hence it is possible that for some parameter configurations less than 15% of the data are in regime 1 or 2. For  $\varsigma$  we assume a Bernoulli prior with prior hyperparameter  $\xi$ .

The previous priors do not require any serious subjective prior input. In our empirical exercise, we are interested in comparing hypotheses about the presence of nonlinearity and heteroskedasticity. Hence, informative priors over all the remaining parameters are required. For  $\beta$  we use a  $N(d_0, D_0)$  prior. We assume the error variances are, *a priori*, independent of each other and that  $\sigma_m$  is inverted Gamma with hyperparameters  $\underline{\mu}_m$  and  $\underline{\nu}_m$  (we adopt the notation for the inverted-Gamma given in Judge, Griffiths, Hill, Lutkepohl and Lee (1985), pages 106-107). Prior degrees of freedom are given by  $\underline{\nu}_m$  and the prior mode of  $\sigma_m$  is  $\sqrt{\frac{\underline{\nu}_m}{\underline{\nu}_m+1}}\underline{\mu}_m$ . We assume  $\beta$ ,  $\sigma$  and  $\gamma$  are all *a priori* independent of each other.

In our empirical analysis of the sunspot data, we set  $d_0$  to zero, and  $D_0 = .45I$ . For the error variances we set  $\underline{\nu}_m = 3$  and  $\underline{\mu}_m = 10$ . We set  $\xi = 0.5$  and  $\bar{K} = \bar{J} = p$ . All these values are relatively noninformative.

### 3.3 Markov Chain Monte Carlo

If  $\gamma$  were known, the model reduces to a heteroskedastic linear regression model. Hence, conditional on  $\gamma$ , a posterior simulator can be set up in a straightforward manner. The posterior for  $\gamma$ , unfortunately, is of a more complicated form and, hence, we use a Metropolis-Hastings algorithm.

The analysis is conducted in terms of the conditional likelihood with the first  $\max(p, \bar{K} + \bar{J})$  values of  $Y_t$  assumed fixed. Further, the pre-sample values of  $I_{1t}, I_{2t}$  are set to zero. The log posterior density of  $\psi$  is given by:

$$\begin{aligned} \ln P[\psi|\mathbf{Y}] &= \ln P[\beta, \sigma, r_1, r_2, K_1, K_2, J_1, J_2, \varsigma|\mathbf{Y}] \propto \\ & - \sum_{t=1}^N \ln h_t(\gamma, \sigma) - 1/2 \sum_{t=1}^N \frac{(Y_t - \beta' \mathbf{W}_t(\gamma))^2}{h_t(\gamma, \sigma)^2} - \\ & - \sum_{m=0}^2 \left( \frac{\underline{\nu}_m}{2} - 1 \right) \ln(\sigma_m) - |D_0|^{1/2} - (\beta - d_0)' D_0^{-1} (\beta - d_0) + \ln(f(r_2|\kappa, r_1)f(r_1|\kappa)) + \varsigma \ln(\xi) + (1 - \varsigma) \ln(1 - \xi), \end{aligned} \quad (18)$$

on the support of  $\Gamma$ .

The MCMC algorithm can be set up in terms of the following conditional densities:

1. Conditional on  $(\sigma, \gamma, \mathbf{Y})$  the distribution of  $\beta$  is multivariate normal with mean:

$$E[\beta|\sigma, \gamma, \mathbf{Y}] = \left\{ D_0^{-1} + \sum_{t=1}^N \mathbf{W}_t^*(\gamma) \mathbf{W}_t^{*'}(\gamma) \right\} \left\{ D_0^{-1} d_0 + \sum_{t=1}^N \mathbf{W}_t^{*'}(\gamma) Y_t^*(\gamma) \right\}$$

where

$$\mathbf{W}_t^*(\gamma) = \mathbf{1}[I_{0t-1}(\gamma) = 1] \frac{\mathbf{W}_t(\gamma)}{\sigma_0} + \mathbf{1}[I_{1t-1}(\gamma) > 0] \frac{\mathbf{W}_t(\gamma)}{\sigma_1} + \mathbf{1}[I_{2t-1}(\gamma) > 0] \frac{\mathbf{W}_t(\gamma)}{\sigma_2},$$



$$Y_t^*(\gamma) = \mathbf{1}[I_{0t-1}(\gamma) = 1] \frac{Y_t}{\sigma_0} + \mathbf{1}[I_{1t-1}(\gamma) > 0] \frac{Y_t}{\sigma_1} + \mathbf{1}[I_{2t-1}(\gamma) > 0] \frac{Y_t}{\sigma_2},$$

and covariance:

$$\left\{ D_0^{-1} + \begin{bmatrix} \Sigma_{11} & \sigma_1^{-2} \sum_{t=1}^N Z_{1t-1}(\gamma) \mathbf{X}_t & \sigma_2^{-2} \sum_{t=1}^N Z_{2t-1}(\gamma) \mathbf{X}_t \\ \sigma_1^{-2} \sum_{t=1}^N Z_{1t-1}(\gamma) \mathbf{X}_t' & \sigma_1^{-2} \sum_{t=1}^N Z_{1t-1}^2(\gamma) & 0 \\ \sigma_2^{-2} \sum_{t=1}^N Z_{2t-1}(\gamma) \mathbf{X}_t' & 0 & \sigma_2^{-2} \sum_{t=1}^N Z_{2t-1}^2(\gamma) \end{bmatrix} \right\}^{-1}, \quad (19)$$

with the  $(p+1) \times (p+1)$  sub-matrix,

$$\Sigma_{11} = \sigma_0^{-2} \sum_{t=1}^N \mathbf{1}[I_{0t-1}(\gamma) = 1] \mathbf{X}_t \mathbf{X}_t' + \sigma_1^{-2} \sum_{t=1}^N \mathbf{1}[I_{1t-1}(\gamma) > 0] \mathbf{X}_t \mathbf{X}_t' + \sigma_2^{-2} \sum_{t=1}^N \mathbf{1}[I_{2t-1}(\gamma) > 0] \mathbf{X}_t \mathbf{X}_t'.$$

Thus standard algorithms for generating draws from a multivariate normal can be used.

2. Conditional on  $(\{\sigma_\ell : \ell \neq i\}, \gamma, \beta, \kappa, \mathbf{Y})$  the distribution of  $\sigma_m$  is inverted Gamma with parameters  $\bar{\nu}_m$  and  $\bar{\mu}_m$ . The posterior degrees of freedom and mean can be obtained using  $\bar{\nu}_m = \underline{\nu}_m + \sum_{t=1}^N \mathbf{1}[I_{mt}(\gamma) > 0]$  and  $\bar{\mu}_m^2 = \frac{\underline{\nu}_m \underline{\mu}_m^2 + \sum_{t=1}^N e_{mt}^2}{\bar{\nu}_m}$  where  $e_{mt} = \mathbf{1}[I_{mt}(\gamma) > 0](Y_t^* - \beta' \mathbf{W}_t^*)$ . Thus standard algorithms for generating draws from an inverted Gamma distribution can be used.
3. Conditional on  $(\beta, \sigma, \mathbf{Y})$  we split  $\gamma$  into three main groups,  $(r_1, r_2)$ ,  $(\kappa)$  and  $\varsigma$  and use the notation  $\gamma_{-r_m}$  to indicate the vector  $\gamma$  with  $r_m$  omitted.

- (a) In order to draw the thresholds we use a Metropolis algorithm as in Chen and Lee (1995). Of course such a simple algorithm will introduce a great deal of dependence in the draws from the posterior but this drawback is compensated for by its speed of calculation compared to other methods. Two threshold draws are required, hence first we condition on  $r_2$ . The proposed threshold value  $r_1'$  is drawn from a uniform with lower support fixed by the value of  $r_2$  and the requirement that at least 15% of the data be in major regime 0 and upper support by the requirement that at least 15% of the data be in major regime 1. If

$$\left[ - \sum_{t=1}^N \ln h_t(r_1', \gamma_{-r_1}, \sigma) - 1/2 \sum_{t=1}^N \frac{(Y_t - \beta' \mathbf{W}_t(r_1', \gamma_{-r_1}))^2}{h_t(r_1', \gamma_{-r_1}, \sigma)^2} \right] -$$

$$\left[ - \sum_{t=1}^N \ln h_t(r_1, \gamma_{-r_1}, \sigma) - 1/2 \sum_{t=1}^N \frac{(Y_t - \beta' \mathbf{W}_t(r_1, \gamma_{-r_1}))^2}{h_t(r_1, \gamma_{-r_1}, \sigma)^2} \right]$$

is larger than the logarithm of a draw from a standard uniform then the value of the  $r_1$  is changed to  $r_1'$ , otherwise the present value is retained. This process is then repeated for  $r_2$  conditioning on  $r_1$ .

- (b) Since,  $K_1, J_1, K_2, J_2$  only take on a discrete number of values, we can evaluate the conditional posterior density at all possibilities to produce a discrete distribution to draw from. However, this requires  $\bar{J}^2 \times (1 + \bar{K})^2$  evaluations of the likelihood function. Each evaluation requiring the creation of the  $\{I_{mt}, Z_{mt}, m = 1, 2\}$  variables. Thus, in practice it is computationally much faster to use a Metropolis algorithm. Hence we draw a proposed quadruple  $\kappa' = (K'_1, J'_1, K'_2, J'_2)$  from the joint prior distribution over these parameters. Again if

$$\left[ -\sum_{t=1}^N \ln h_t(\kappa', \gamma_{-\kappa}, \sigma) - 1/2 \sum_{t=1}^N \frac{(Y_t - \beta' \mathbf{W}_t(\kappa', \gamma_{-\kappa'}))^2}{h_t(\kappa', \gamma_{-\kappa} \sigma)^2} + \ln(f(r_2|\kappa', r_1)f(r_1|\kappa')) \right] -$$

$$\left[ -\sum_{t=1}^N \ln h_t(\kappa, \gamma_{-\kappa}, \sigma) - 1/2 \sum_{t=1}^N \frac{(Y_t - \beta' \mathbf{W}_t(\kappa_1, \gamma_{-\kappa}))^2}{h_t(\kappa, \gamma_{-\kappa} \sigma)^2} + \ln(f(r_2|\kappa, r_1)f(r_1|\kappa)) \right]$$

is larger than the logarithm of a draw from a standard uniform then the value of the  $\kappa$  is updated to  $\kappa'$ , otherwise the present value is retained.

- (c) For  $\varsigma$  we evaluate the posterior density at the two possible values to construct a Bernoulli distribution to draw from.

For all data sets and models, we take 20,000 MCMC replications and discard the initial 2,000 to minimize start-up effects.

### 3.4 Bayes Factors

Koop and Potter (1999a) discuss the use of Bayes factors in nonlinear time series models and the reader is referred to that paper for a detailed discussion. Here, we briefly note that a common criticism of nonlinear time series models is that they are parameter rich. In linear time series analysis a standard means of penalizing over-parameterization is to use information measures that mimic the Occam's razor property of Bayes factors. In the nonlinear time series case, there is less justification for information measures approximating Bayes factors because of the potential for multiple peaks in the likelihood function and because of difficulties in deciding the penalty for the regime classification parameters given their fast convergence. Thus, in this paper we rely directly on Bayes factors to penalize parameter rich models. The general formula for the Bayes factor for model 1 parameterized by  $\zeta$  versus model 2 parameterized by  $\psi$  is given by the ratio of the marginal likelihoods:

$$B_{1vs.2} = \frac{\int \ell(\zeta)p(\zeta) d\zeta}{\int \ell(\psi)p(\psi) d\psi},$$

where  $p(\zeta), p(\psi)$  are the (proper) prior distributions over  $\zeta$  and  $\psi$  respectively and  $\ell(\zeta)$  and  $\ell(\psi)$  are likelihood functions.

In order to calculate Bayes factors for the EDTAR model versus linear models from the output of the Markov Chain Monte Carlo algorithm we use the approach of Koop and Potter (1999a). Note that, under the restriction  $\theta_1 = \theta_2 = 0$ , the nonlinearity in the conditional mean of the EDTAR model disappears and one is left with a conditionally heteroskedastic linear model:

$$Y_t = \phi' \mathbf{X}_t + h_t(\gamma, \sigma) V_t.$$

We will call this the LEDTAR model. The LEDTAR model is nested within the EDTAR model and under the additional assumption that the priors of the two models over shared parameters are equal at  $\theta_1 = \theta_2 = 0$ , one can calculate the Bayes factor using the Savage-Dickey density ratio:

$$B_{\text{Linear vs EDTAR}} = \frac{p(\theta_1 = 0, \theta_2 = 0 | \mathbf{Y})}{p(\theta_1 = 0, \theta_2 = 0)}.$$

That is, the Bayes factor is given by the ratio of the height of the posterior density at  $\theta_1 = \theta_2 = 0$  to the prior density height at this point. Under our normal prior assumption, the denominator of this expression is easy to calculate. The numerator can be found by averaging the Normal conditional density:

$$p(\theta_1 = 0, \theta_2 = 0 | \gamma, \sigma, \mathbf{Y}),$$

across draws of  $\gamma, \sigma$  from the MCMC algorithm.

Further, under the restriction that  $\sigma_1 = \sigma_2 = \sigma_0$ , the nonlinearity in the conditional variance is no longer relevant. Using the Savage-Dickey density ratio, again under the restriction on the priors of the shared parameters, one can calculate the Bayes factor of a homoskedastic autoregressive model against the LEDTAR (see the appendix to Koop and Potter (1999a)). However, in order to do this one must run a separate MCMC algorithm for the LEDTAR model.

Multiplying these two Bayes factors together produces a Bayes factor for the EDTAR model against a linear AR model with same prior over shared parameters. It is also possible to calculate a Bayes factor for a homoskedastic EDTAR model, i.e.,  $\sigma_0 = \sigma_1 = \sigma_2$ , versus a heteroskedastic EDTAR model during the MCMC run for the heteroskedastic EDTAR (i.e. simply use the same technique as for the LEDTAR versus homoskedastic linear autoregressive model).

The estimates of the marginal likelihoods are all conditional on  $p$ , the number of autoregressive lags. Thus, by repeating the analysis for different values of  $p$ , one can obtain a posterior distribution across  $p$  for each type of model.

## 4 Empirical Illustrations

### 4.1 Sunspot Data

In this section we illustrate the use of the EDTAR model on sunspot data obtained from Tong (1990). This data set includes annual observations on sunspot activity from 1700 to 1988. It is widely thought that this series is nonlinear, but there is no consensus as to which nonlinear model is appropriate. This data set has been analyzed using two regime SETAR models by many authors such as Tong (1990), Geweke and Terui (1993) and Chen, McCulloch and Tsay (1996). We work with demeaned data.

We compare three classes of models: the linear autoregression,  $LAR(p)$ , the two regime threshold model with delay variable defined off the level of the lag of the sunspot data,  $SETAR(p)$  and the three regime EDTAR( $p$ ). With the two latter classes, we consider a homoskedastic and heteroskedastic version where the heteroskedasticity is modeled as different error variances for the major regimes. In all cases, the maximum value of  $p$  we consider is 11. This value is roughly equal to the length of the cycle in the data. Previous work with this data have tended to find  $p = 3$  or 9 in  $LAR(p)$  or  $SETAR(p)$  models so a choice of  $p = 11$  should nest all relevant models. The delay parameter,  $d$ , in the SETAR model is treated as an unknown parameter and its prior is flat over the integers  $1, 2, \dots, p$ . The prior for the threshold parameter for this model is Uniform over values of the delay which imply at least 15% of the data in each regime. All models have an intercept; for the SETAR( $p$ ) this intercept varies across regimes. For all models, we treat initial conditions as fixed and all errors are Normally distributed.

For the autoregressive coefficients and error variance(s) in the LAR and TAR models, we use the standard natural conjugate prior with prior hyperparameters selected so as to match the EDTAR prior. That is, for the autoregressive coefficients in both models, we use a prior mean of 0 and prior covariance of  $.45I$  and, for the error standard deviation(s), we set prior degrees of freedom equal to 3 and the mean hyperparameter,  $\underline{\mu}_m = 10$ . Note that the priors used for the LAR and SETAR are not identical to the EDTAR since the natural conjugate prior used in the former models does not have prior independence between the autoregressive coefficients and error variance(s). In order to compare the EDTAR to SETAR and LAR we need to calculate the marginal likelihood for the LAR( $p$ ) model with independent Normal-Gamma prior. We do this using the method of Chib (1995).

Under these assumptions, the LAR( $p$ ) model is in the form of a linear regression model

with natural conjugate prior and textbook results can be used to calculate the marginal likelihood. The SETAR( $p$ ) is defined in Equation (2). Calculation of marginal likelihoods is straightforward since, conditional on knowing the threshold, the model breaks down into two linear regression models. Hence, conditional on the threshold, standard results for the linear regression model with natural conjugate priors can be used to calculate the marginal likelihood. To get an unconditional marginal likelihood, the threshold is integrated out in the usual way. Since the threshold can be treated as a discrete variable, this integration step is easy (see Koop and Potter, 1999b, for more details).

Table 1 presents posterior model probabilities (to 4 decimal places) for all models for all choices of  $p$ . Reading down a column reveals evidence regarding the choice of  $p$  for each model type. For the LAR and SETAR models  $p = 9$  receives the highest probability. For the AR model, if one were to restrict  $p$  to be relatively small (e.g. less than 6),  $p = 3$  would be chosen. These findings of evidence of  $p$  being 3 or 9 accords with previous studies. The SETAR outperforms the LAR model, although there is little evidence of heteroskedasticity based on the delay variable used by the SETAR. However, the key finding of this paper is that the EDTAR with  $p = 3$  receives most of the posterior probability. For the EDTAR there is strong evidence for heteroskedasticity. It is interesting to note that the tightly parameterized three regime EDTAR( $p = 3$ ) with 15 parameters does better than models such as the two regime SETAR( $p = 9$ ) with 23 parameters.

**Table 1: Estimated Posterior Model Probabilities**

$p$	LAR	SETAR-homo	SETAR-het	LEDTAR	EDTAR-hom	EDTAR-het
1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2.0000	0.0000	0.0000	0.0000	0.0042	0.0000	0.0007
3.0000	0.0000	0.0000	0.0000	0.1751	0.0000	0.7548
4.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001
5.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
7.0000	0.0000	0.0003	0.0000	0.0002	0.0000	0.0000
8.0000	0.0000	0.0007	0.0000	0.0000	0.0000	0.0000
9.0000	0.0000	0.0237	0.0000	0.0013	0.0000	0.0001
10.0000	0.0000	0.0228	0.0000	0.0011	0.0000	0.0005
11.0000	0.0000	0.0143	0.0000	0.0000	0.0000	0.0000

Table 2 presents the posterior distributions of the discrete variables for the EDTAR-het model with  $p = 3$ . There is strong evidence that  $K_1$  and  $K_2$  are both one, indicating that last year's changes in sunspot activity trigger the changes in major regime. Our results indicate that  $J_1$  and  $J_2$  are 1, indicating that only one activation is required to trigger a regime shift. Thus,

the regimes identified are of upswings and downswings in the sunspot numbers rather than from the absolute level of the numbers as in the SETAR model. The parameter  $\varsigma$  is 0 with probability .5. That is, its prior and posterior are the same. This occurs because, for this particular data set and the realized posterior values of  $\{K_m, J_m, r_m, m = 1, 2\}$ , there is never a time where both regime 1 and regime 2 could be simultaneously active.

**Table 2: Posterior Distribution of  $\kappa$**

Lag	$K_1$	$K_2$	$J_1$	$J_2$
0	0.001	0.000	n.a.	n.a.
1	0.998	0.986	1.000	0.979
2	0.001	0.012	0.000	0.020
3	0.000	0.002	0.000	0.001

Table 3 presents posterior means and standard deviations for all of the continuous parameters of the EDTAR-het( $p = 3$ ) model. The posterior means of  $\theta_1$  and  $\theta_2$  are both negative and are more than three standard deviations from zero. Hence, a strong dampening is observed, especially for the downswing regime. For the upswing regime, there is strong evidence that the error variance is much larger than the other regimes. These findings account for the strong results in Table 1. The posterior means for the  $\sigma'_m$ s differ across regimes, which accounts for the evidence of heteroskedasticity. The thresholds defining the major regimes are quite precisely estimated with the upswing regime requiring a smaller change for its activation.

**Table 3: Posterior Properties Of EDTAR-het**

Parameter	Posterior Mean	Posterior Standard Deviation	Convergence Diagnostic
$\alpha$	-1.5	1.2	1.618
$\phi_1$	1.332	0.101	-1.204
$\phi_2$	-0.410	0.109	0.374
$\phi_3$	-0.247	0.049	-0.618
$\theta_1$	-0.103	0.030	0.732
$\theta_2$	-0.461	0.137	0.639
$r_1$	10.6	1.8	—
$r_2$	-17.5	1.6	—
$\sigma_0$	12.4	0.1	-0.089
$\sigma_1$	25.7	2.3	1.374
$\sigma_2$	10.9	1.1	0.698

The last column of Table 3, presents the convergence diagnostic suggested by Geweke (1992) for the continuous parameters of the model. It compares the difference in estimated posterior means between the first and last 10% of the MCMC replications relative to the numerical

standard error. As the number of MCMC replications increases, this diagnostic should have a  $N(0,1)$  distribution. It can be seen that convergence does seem to have occurred. Informal experimentation with different starting values and numbers of replications also indicate that convergence is not a problem for our MCMC algorithm.

$R^2$  is the standard measure of fit used in regression models. Here we define a Bayesian analogue as:

$$R^2 = 1 - \frac{\sum e_{0t}^2 + \sum e_{1t}^2 + \sum e_{2t}^2}{\sum (Y_t - \bar{Y})^2},$$

where  $e_{jt}$  has been defined previously as the errors in regime  $j$ . Using the MCMC parameter draws, we can integrate over the parameter space and calculate  $E(R^2)$ . For the EDTAR-het with  $p = 3$  we find  $E(R^2) = 0.828$ .

## 4.2 Artificial Data

To supplement empirical work with real data, it is always instructive to consider artificial data in order to provide additional support for the contention that our Bayesian methods are working well and can distinguish between relevant model classes. Given that Bayesian analysis of the EDTAR model requires posterior simulation, it would be computationally burdensome to carry out a full Monte Carlo study. Accordingly, we simulate three artificial data sets, one from a linear AR, one from an EDTAR and one from another nonlinear time series model. In particular, we generate  $T = 250$  observations, assuming  $IIIDN(0,1)$  errors (i.e. no heteroskedasticity in any of the models), from the following models:

- An AR(1) model with intercept = 0.4 and AR(1) coefficient = 0.4.
- An EDTAR model with AR(1) structure as above in the regime defined by  $I_{0t} > 0$ . In addition, we set  $K_1 = K_2 = 0$ ,  $J_1 = J_2 = 1$ ,  $r_1 = -0.5$ ,  $r_2 = 0.5$  and  $\theta_1 = \theta_2 = -0.25$ .
- A two regime SETAR model. If the lagged dependent variable is less than a threshold value of 0.0, then an AR(1) model with intercept = 0.8 and AR(1) coefficient = 0.8 applies. Else, an AR(1) model with intercept = 1.0 and AR(1) coefficient = 0.0 applies.

For simplicity, we will focus on the EDTAR and AR models (homoskedastic and heteroskedastic versions) and restrict attention to the case  $p = 1$ . We use the same prior as for the sunspot data, except for  $\sigma_m$ . For these, the prior degrees of freedom remain at 3, but the

prior mean parameters for the  $\{\sigma_m, m = 0, 1, 2\}$  are changed to 1.0. Assuming equal prior odds, we can calculate the posterior model probabilities for each of our four models for the three different artificial data sets. There are given in Table 4.

**Table 4: Posterior Model Probabilities for Different Data Generating Processes**

DGP\Model	LAR	LEDTAR	EDTAR-hom	EDTAR-het
AR(1)	0.8091	0.0497	0.1344	0.0068
EDTAR	0.0211	0.0025	0.8977	0.0787
TAR	0.3464	0.0295	0.5655	0.0586

It can be seen that the Bayes factors indicate the correct model with high probability. When the true data generating process is AR(1), the correct LAR model receives 81% of the posterior model probability. When the DGP is EDTAR, the correct EDTAR-hom model receives 90% of the posterior model probability. Note also that the Bayes factors correctly allocate little weight to heteroskedastic models. The last row of Table 4 considers the interesting case where the true DGP is a nonlinear time series process other than the EDTAR. This row provides some evidence that nonlinearity is present, but the linear model also receives appreciable support. In other words, the last row indicates, reasonably, a good deal of uncertainty over which model is correct.

Using the unrestricted EDTAR model, posterior means (not reported here) are close to the true values used to generate the data (relative to posterior standard deviations) for the first two artificial data sets. Since the third artificial data set is not nested in the EDTAR, a similar comparison of EDTAR posterior means to true values is not possible.

## 5 Conclusions

In this paper, we have developed Bayesian methods to examine the posterior properties of a general class of EDTAR models which is more flexible than that previously used in the literature. We argue that this class is valuable as parsimonious threshold models which capture the types of behavior often observed in time series data. The empirical example indicates that, on the often used sunspot data, the EDTAR outperforms both linear and traditional threshold models.

Two extensions of EDTAR models are of interest. First, in the development of parameter restriction across sub-regimes in section 2.3, we imposed the very tight restriction that  $\xi_{ms} = \theta_m$ . Instead, one could consider approximating the moving average weights  $\xi_{ms}, s = 0, 1, \dots$  by ratios of low order polynomials in the lag operator. The extension of the MCMC algorithm presented above



to this case is straightforward and Bayes factors can be used to evaluate the value of relaxing the restriction.

Second, EDTAR models and the Bayesian methods used here are particularly well-suited to the major task of nonlinear modeling of multiple time series. A major difficulty with standard SETAR modeling in this case is the vast number of autoregressive parameters that are required. For example, in a four variable system with maximum autoregressive lag of 5 (a common situation in economics) and three regimes the number of autoregressive parameters would be 240. In Koop, Pesaran and Potter (1996), the EDTAR model is extended to the multiple time series case by using endogenous delay variables defined by a single variable but generating different  $Z_{mt}$  variables for each time series in the system. The collection of  $Z_{mt}$  variables for each regime are then allowed to enter the equation for each variables. In a four variable system with  $p = 5$  this requires the estimation of only 112 autoregressive type parameters.

## References

- [1] Beaudry, P. and Koop, G. (1993). "Do recessions permanently change output?" *Journal of Monetary Economics*, 31, 149-163.
- [2] Chan, K.S. (1993). "Consistency and Limiting Distribution of a Least Squares Estimator of a Threshold Autoregressive Model," *Annals of Statistics*, 31, 520-533.
- [3] Chen, R., (1990). "Threshold Variable Selection in Open-loop Autoregressive Models," *Journal of Time Series Analysis* 16, 461-482.
- [4] Chen, C.W.S. and Lee, J.C. (1995) "Bayesian Inference of Threshold Autoregressions," *Journal of Time Series Analysis* 16, 483-492.
- [5] Chen, C.W.S , McCulloch, R. and Tsay, R., (1996) "A Unified Approach to Estimation and Modeling of Linear and Nonlinear Time Series," mimeo, University of Chicago.
- [6] Chib, S., (1995). "Marginal Likelihood from the Gibbs output," *Journal of the American Statistical Association*, 90, 1313-1321.
- [7] Gelfand, A.E. and A.F.M. Smith (1990). "Sampling-based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, 398-409.

- [8] Geweke, J. (1992). "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," pp. 169-193 in Bayesian Statistics 4, eds. J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith, Oxford: Oxford University Press.
- [9] Geweke, J. and N. Terui (1993). "Bayesian Threshold Autoregressive Models for Nonlinear Time Series," Journal of Times Series Analysis, 14, 441-454.
- [10] Hamilton, J. (1989). "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," Econometrica, 57, 357-384.
- [11] Hess, G. and Iwata, S. (1997). "Asymmetric Persistence in GDP? A Deeper Look at Depth," Journal of Monetary Economics, 40, 535-554.
- [12] Jansen, D. and Oh, W. (1999). "Modeling Nonlinearity of Business Cycles: Choosing Between the CDR and STAR Models," Review of Economics and Statistics, 81, 344-349.
- [13] Judge, G., Griffiths, W., Hill, R.C. and Lee, T.-C. (1985). The Theory and Practice of Econometrics, second edition, New York: John Wiley and Sons.
- [14] Koop, G., Pesaran, M.H. and Potter, S., (1996). "Impulse Response Analysis in Nonlinear Multivariate Models," Journal of Econometrics, 74, 119-147.
- [15] Koop, G. and Potter, S. (1999a). "Bayes Factors and Nonlinearity: Evidence from Economic Time Series," Journal of Econometrics, 88, 251-281.
- [16] Koop, G. and Potter, S. (1999b). "Dynamic Asymmetries in US Unemployment," Journal of Business and Economic Statistics, 17, 298-312.
- [17] Pesaran, M.H. and Potter, S. (1997). "A Floor and Ceiling Model of U.S. Output," Journal of Economic Dynamics and Control, 21, 661-695.
- [18] Potter, S. (1995). "A nonlinear approach to US GNP," Journal of Applied Econometrics, 10, 109-126.
- [19] Tiao, G. and Tsay, R. (1994). "Some Advances in Non-linear and Adaptive Modelling in Time Series," Journal of Forecasting, 13, 109-131.
- [20] Tong, H. (1990). *Non-linear Time Series Models*, Clarendon Press, Oxford.

- [21] Tong, H. and Lim, K. (1980). “Threshold autoregressions, limit cycles and cyclical data,”  
Journal of the Royal Statistical Society Series B, 42, 245-292.